Presentation originally made on August 6, 2008 at the contributed section

This version is a revision of October 2008 based on feedback from various sources

"Counting the Dead in Iraq"

By

Michael Spagat
Royal Holloway College
University of London

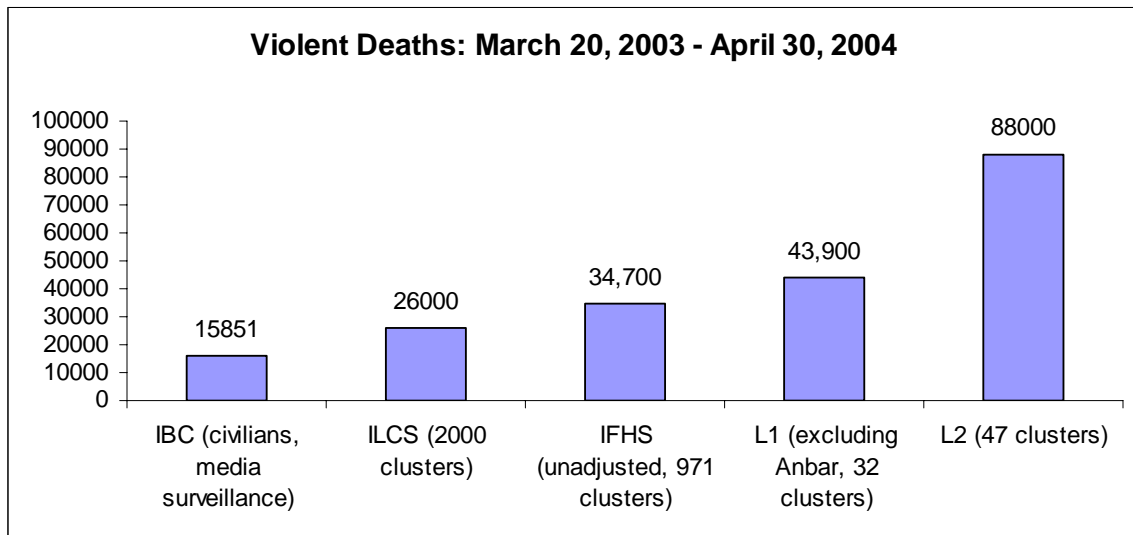JSM2008
Denver, Colorado

Main Sources on Mortality in Iraq

Roberts et al. (2004) – L1

Iraq Living Conditions Survey 2004 – ILCS

Burnham et al. (2006) – L2

Iraq Family Health Survey (2008) – IFHS

Iraq Body Count project (ongoing) - IBC

Violent Deaths: March 20, 2003 - April 30, 2004

Source	Value
IBC (civilians, media surveillance)	15851
ILCS (2000 clusters)	26000
IFHS (unadjusted, 971 clusters)	34,700
L1 (excluding Anbar, 32 clusters)	43,900
L2 (47 clusters)	88000

L2's central estimate for violent deaths for the first 13 months of the war exceeds:

1. The media-based count of civilian deaths by IBC by **more than a factor of 5;**

2. The ILCS central estimate by **more than a factor of 3;**

3. The IFHS central estimate by **more than a factor of 2.5;**

4. The L1 central estimate **by more than a factor of 2.**

Figure 1 - Violent Deaths:
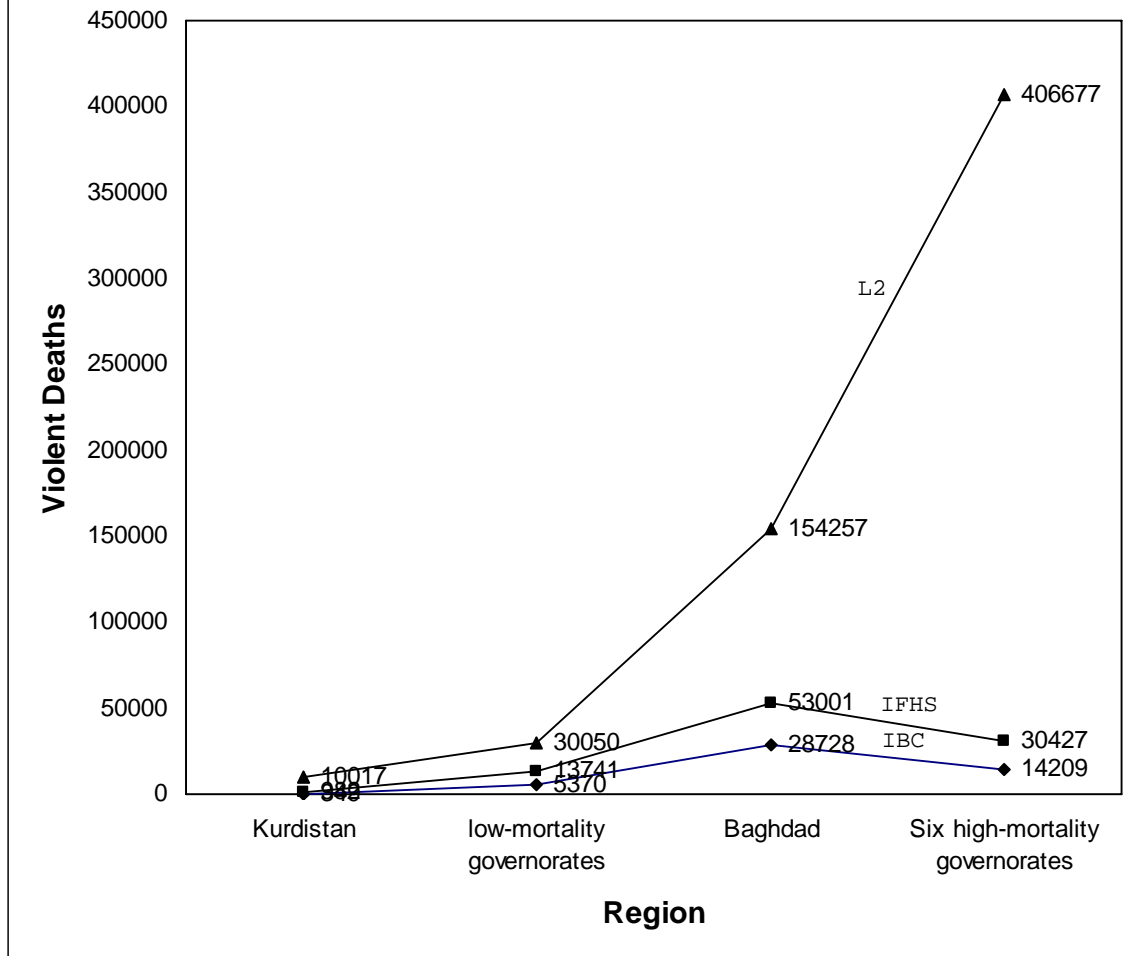**Three sources, three time periods**
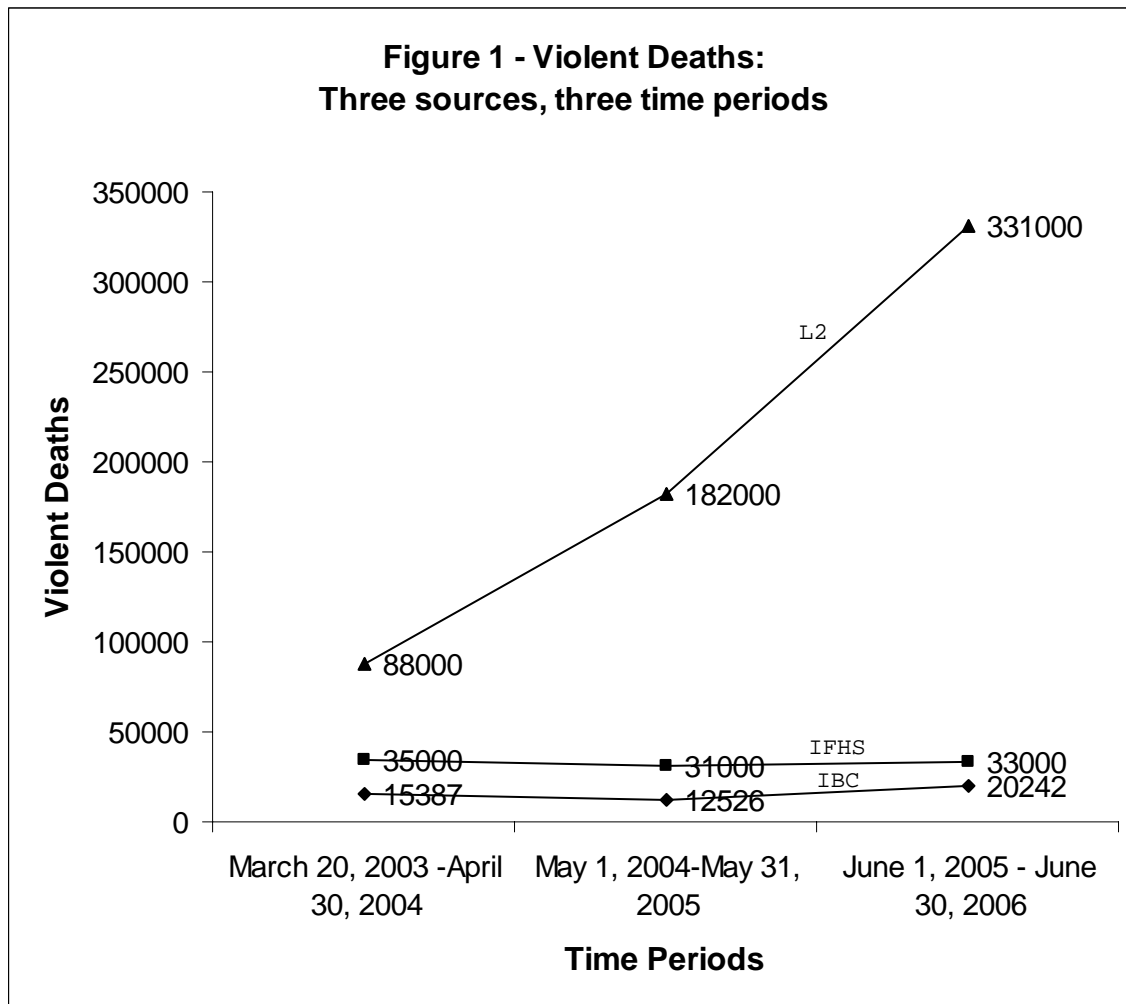
The **gap** between L2 and other sources **grows sharply over time**.

During the second 13 months of the war the L2/IFHS ratio is **nearly 6** and the L2/IBC ratio is **nearly 15**.

During the third 13 months of the war the L2/IFHS ratio is **more than 10** and the L2/IBC ratio is **nearly 16**.

Figure 2 - Regional Breakdown:
March 20, 2003 - July 10, 2006

In six high-mortality governorates L2 exceeds the IFHS by **more than a factor of 13** and the IBC by **nearly a factor of 29.**

The ratios are **2.9 and 5.4 in Baghdad, 2.2 and 5.6 in low-mortality governorates and 10.2 and 28.7 in Kurdistan (respectively).**

Figure 3 - Regional Breakdown:
March 20, 2003 - April 30, 2004
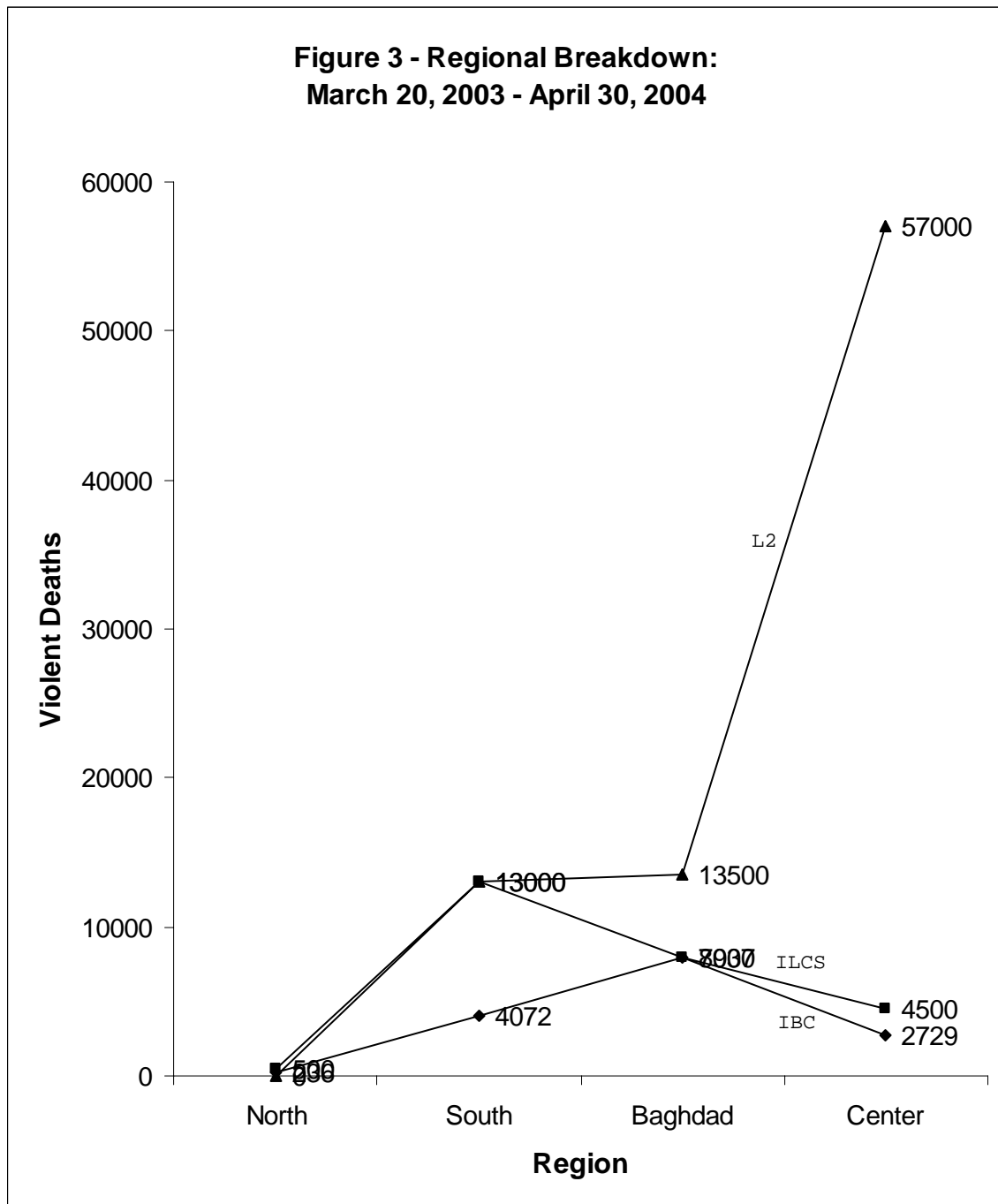
In the center L2 exceeds the ILCS and IBC by factors of 12.7 and 20.9 respectively. In Baghdad these ratios are both 1.7.

Summary

Compared to the other main sources on violent deaths in Iraq L2 shows:

1. **Many more** violent deaths;

2. A **much sharper upward time trend** in violent deaths;

3. A **much different geographical distribution** of violent deaths.

Data from many other sources, including the United Nations Assistance Program for Iraq (UNAMI), the Iraqi Ministry of Health, the Baghdad Morgue and the US Department of Defense, are either also inconsistent with L2 on these points or not relevant (e.g., UNAMI covers only 2006 so is not useful for discerning time trends). No credible data support L2 on any of these three points.

Please show me data if you think otherwise.

If you believe that conflict mortality surveys can make a valuable contribution to human knowledge and/or humanitarian action then you cannot ignore these large discrepancies.

High-quality conflict surveys are expensive, e.g., the ILCS cost more than $1,000,000.

Research funds are scarce.

We need to understand what went wrong with some of these surveys and how to avoid these missteps in the future.

Otherwise, why bother with future (expensive) conflict surveys?

Death Certificates

L2 authors claim that a high percentage of their violent deaths were confirmed by death certificates (70%) so they are credible. But **the pattern of confirmation through death certificates is suspicious:**

1. In a seven-governorate area L2 reports a **perfect record** of death-certificate confirmation when checked for **violent deaths** - 77 out of 77 - in the midst of an **imperfect confirmation record** for **non-violent deaths** - 111 out of 132 - (84%).

2. Outside Nineveh governorate L2 reports **180 confirmations in 180 attempts for violent deaths**. This is not credible compared to either the 80% confirmation rate of L1 or even L2's 92% confirmation rate for non-violent deaths.

3. All 22 failures to confirm were concentrated in the governorate of Nineveh.

24 out of 66 deaths in a particular car bombing appear in the supposedly random L2 sample, spread across 18 separate households within a string of 40 contiguous households.

The bombing ripped through a crowded marketplace and did not damage a row of houses, eliminating a potential explanation for the anomaly.

People living all over the place would have wandered around the marketplace, shuffling themselves randomly.

Did 24 neighbors all set out on a shopping excursion together and mill about the marketplace all holding hands?

L2 reports that interviewers **did not try to confirm any of these 24 deaths through death certificates**.

Cluster 34 (highlights)

There were **at least 18 deaths from air strikes in at least 5 incidents affecting 7 households with no death-certificate confirmation for any of these.** IBC has no air strikes in this governorate in any of these months.

At least 5 deaths in at least 4 separate bombings (at least 3 car bombings) affecting 5 separate households. **What are the odds that victims of at least 3 separate car bombings would live so close together?** IBC does not have a car bomb in Nineveh for the month of 1 of these car-bomb deaths.

Only 9 out or 35 violent deaths in this cluster are reported to be confirmed by death certificate and it contains 19 out of the total of 22 cases where L2 reports that respondents failed to produce death certificates when asked.

Non-disclosure

a. Exact wordings of questions as asked in the field, i.e., a questionnaire.

b. Details of the sample design – The authors have issued a long string of contradictory statements on how the sampling has been done.

c. Data matching annonomized interviewer IDs with survey results.

The last is crucial. There are strange patterns in clusters 33 and 34 and extraordinary numbers of violent deaths in a number of clusters in central governorates surrounding Baghdad. Knowing which of the two interview teams did those clusters would be very valuable information. The L2 authors have not disclosed this information and have even said that **their own field teams did not disclose this information to the authors themselves.**

No contact and failed cluster visits

L2's 0.9% no-contact rate is much lower than the 3.4% no-contact rate for the IFHS. Using the IFHS success rate of 96.6% and assuming statistical independence among household visits then **the odds against the L2 no-contact rate are more than 500,000 to 1 against.**

L2's reports just 1 failed cluster visit due to security problems. The IFHS reports a number of security-related failed cluster visits, including 31 out of 96 in Baghdad and 71 out of 108 in Anbar. Using the IFHS success rate of 67.7% and assuming independence among cluster visit **the odds against L2's 12 successes in 12 attempts in Baghdad are 108 to 1 against and the odds against 5 successes in 5 attempts in Anbar are 214 to 1 against.**

Note the connection with the already suspicious cluster 33 which was in Baghdad.

Extreme Time Pressure

There were 52 clusters of 40 households, 52 days of field work and 2 teams that had to travel all over the country to do the work.

Single teams would break into 2 sub-teams and complete 40 interviews in a day, lasting 20 minutes on average (it is claimed). Variance around the average would raise the total time spent in some clusters.

First, teams had to complete the final stage of sampling: travel the length of a main street listing all the side streets, select one at random, travel the length of this side street listing all its households and select one of these at random for the first interview.

It was terribly hot: breaks were necessary.

Middle Eastern hospitality and the need to listen respectfully to painful stories could slow things down considerably.

Presence of Known Fabrication Risk Factors

I use and [ASA/AAPOR list of fabrication risk factors](#).

Inadequate supervision

The US-based authors **never in Iraq** during the field work.

There was **no use of recontact methods,** either to check that interviews were actually done or to check that interviews were done properly.

There was **no collection of demographic information** and, therefore, no use of such information to check the veracity of collected data.

There was **no analysis of data matching annonomized interviewer IDs with interview results.**

Poor quality control

Data that are sometimes **missing** include **household sizes (13 times), months in which deaths occurred (57 times),** and the **number of males and females in each household (55 times).**

Data on the number of males and females in each household are missing for an entire cluster.

14% of the time there is failure of the identity,

Household size 2006 = Household size 2002 + births – deaths + in migration – out migration

Excessive workload (extreme)

**Offsite isolation of interviewers from the parent organization**

A variety of independent factors all point in the direction of fabrication.

For example, poorly supervised field workers under extreme time pressure reported, improbably, 12 successful cluster visits in 12 attempts in Baghdad. In one of these clusters, Cluster 33, they reported 24 victims of a single car bombing. They did not attempt to verify any of these deaths through death certificates.

In Nineveh, the IFHS failed to cover 10 clusters out of 72. The rushing L2 team succeeded 5 times out of 5 and found strange violence and death-certificate patterns in cluster 34.

Evidence of Extrapolation of the L2
Results from Previous Studies
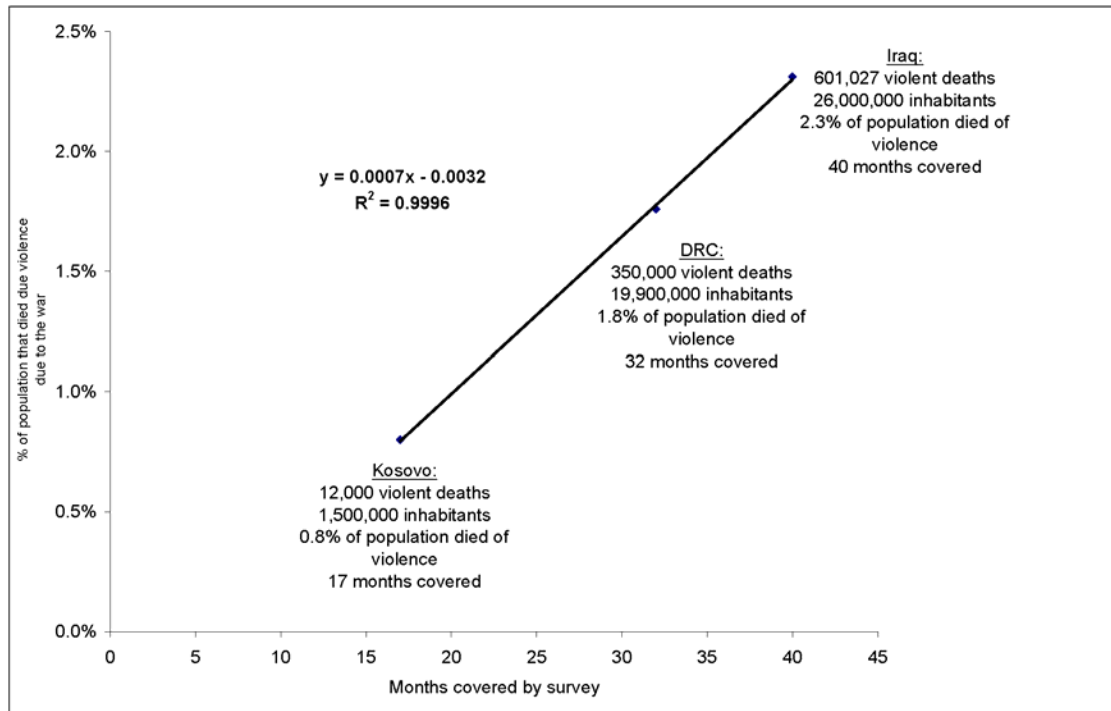
In a letter to the *Lancet* the L2 authors cited a study done by Les Roberts and others in the DRC that found that 1.8% of the population was violently killed in 32 months[1] This is cited as making it plausible that 2.3% of the population in Iraq might have been violently killed in 40 months of conflict as claimed by L2.

The L2 authors cite in L2 and elsewhere a Kosovo finding that 0.8% of the population was killed in 17 months of conflict.

It turns out that these three data points [(17, 0.8), (32, 1.8), (40, 2.3)] are in near-perfect alignment.[2]

Mark van der Laan of Berkeley calculated that **the probability of such alignment is 0.036, i.e., odds against of 28 to 1. This is evidence of falsification, albeit not definitive.**

---

[1] (33 months according to the letter and even the cited paper sometimes says 33 months but the correct figure is 32 months.)

[2] This is robust to reasonable variation. For example, little changes if you replace the true figure of 32 with the L2 authors' figure of 33.

Conclusion

L2 is inconsistent with all credible evidence on mortality in Iraq (more than I could show in a brief presentation).

L2 is methodologically very weak, e.g., it contains a large number of known risk factors for fabrication.

There is also considerable evidence pointing toward fabrication in L2.

We can tell a consistent story on the levels, trends and geographical patterns of violence in Iraq using just the ILCS, IHFS and IBC.

Adding L2 to the picture generates chaos.

Tossing out L2, based on the evidence at hand, creates space for the field of conflict surveys prosper.

If we cannot bring ourselves to toss out L2 then we remain open to future manipulations.

The most disturbing thing for me in this whole discussion is the number of people who are willing to say privately, but not publicly, that they think that something is seriously wrong with L2.

IBC has now documented nearly 100,000 violent deaths of civilians since the beginning of the war.

The IFHS, now more than two years out of date, estimated 151,000 violent deaths in the war.

In short, there has been tremendous carnage in Iraq.

It is an injustice both to the many victims of this war and to the truth to inflate the number of victims.

There seems to be crude political calculus in play according to which one's estimate of the number of Iraqis killed is increasing in the strength of one's dislike of G.W. Bush.

But the truth must be the most important thing and I hope that statisticians will stand firmly with the truth.